PCT WELTORGANISATION FÜR GEISTIGES EIGENTUM
Internationales Büro
INTERNATIONALE ANMELDUNG VERÖFFENTLICHT NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT)

(51) Internationale Patentklassifikation 6 : G06F 15/80 — A1

(11) Internationale Veröffentlichungsnummer: WO 98/10352

(43) Internationales Veröffentlichungsdatum: 12. März 1998 (12.03.98)

(21) Internationales Aktenzeichen: PCT/DE97/01497

(22) Internationales Anmeldedatum: 15. Juli 1997 (15.07.97)

(30) Prioritätsdaten:
196 35 758.6 3. September 1996 (03.09.96) DE

(71) Anmelder *(für alle Bestimmungsstaaten ausser US)*: SIEMENS AKTIENGESELLSCHAFT [DE/DE]; Wittelsbacherplatz 2, D-80333 München (DE).

(72) Erfinder; und
(75) Erfinder/Anmelder *(nur für US)*: NEUNEIER, Ralf [DE/DE]; Utzschneiderstrasse 10, D-80469 München (DE). ZIMMERMANN, Hans-Georg [DE/DE]; Schiffbauerweg 6 A, D-82319 Starnberg (DE).

(81) Bestimmungsstaaten: JP, US, europäisches Patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Veröffentlicht
Mit internationalem Recherchenbericht.

(54) Title: METHOD AND DEVICE FOR COMPUTER ASSISTED GENERATION OF AT LEAST ONE ARTIFICIAL TRAINING DATA VECTOR FOR A NEURONAL NETWORK

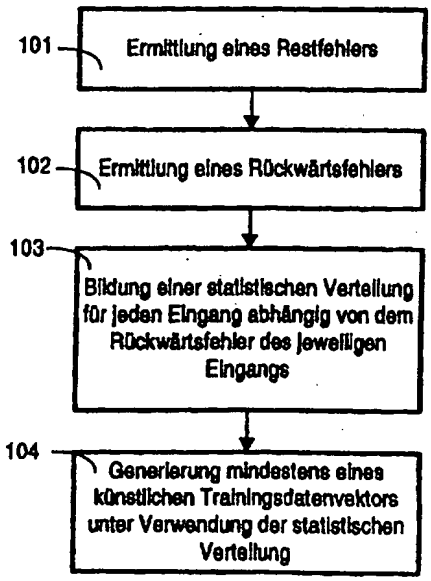
(54) Bezeichnung: VERFAHREN UND VORRICHTUNG ZUR RECHNERGESTÜTZTEN GENERIERUNG MINDESTENS EINES KÜNSTLICHEN TRAININGSDATENVEKTORS FÜR EIN NEURONALES NETZ

(57) Abstract

A residual error (101) is determined after neuronal network training has occurred. A backward error (102) is determined from the residual error. Artificial training data vectors are generated from a statistic random process with an underlying statistic distribution, wherein the respective backward error is considered for a neuronal network input (103, 104).

(57) Zusammenfassung

Nach erfolgtem Training eines neuronalen Netzes wird ein Restfehler ermittelt (101). Aus dem Restfehler wird ein Rückwärtsfehler ermittelt (102). Künstliche Trainingsdatenvektoren werden aus einem statistischen Zufallsprozeß generiert, dem eine statistische Verteilung zugrundeliegt, bei der der jeweilige Rückwärtsfehler für einen Eingang des neuronalen Netzes berücksichtigt wird (103, 104).



101 ... DETERMINATION OF AN UNDETECTED ERROR
102 ... DETRMINATION OF A BACKWARD ERROR
103 ... FORMATION OF A STATISTIC DISTRIBUTION FOR EACH INPUT DEPENDENT ON THE BACKWARD ERROR OF THE RESPECTIVE INPUT
104 ... GENERATION OF AT LEAST ONE ARTIFICIAL TRAINING DATA VECTOR USING STATISTIC DISTRIBUTION

LEDIGLICH ZUR INFORMATION

Codes zur Identifizierung von PCT-Vertragsstaaten auf den Kopfbögen der Schriften, die internationale Anmeldungen gemäss dem PCT veröffentlichen.

AL	Albanien	ES	Spanien	LS	Lesotho	SI	Slowenien
AM	Armenien	FI	Finnland	LT	Litauen	SK	Slowakei
AT	Österreich	FR	Frankreich	LU	Luxemburg	SN	Senegal
AU	Australien	GA	Gabun	LV	Lettland	SZ	Swasiland
AZ	Aserbaidschan	GB	Vereinigtes Königreich	MC	Monaco	TD	Tschad
BA	Bosnien-Herzegowina	GE	Georgien	MD	Republik Moldau	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagaskar	TJ	Tadschikistan
BE	Belgien	GN	Guinea	MK	Die ehemalige jugoslawische Republik Mazedonien	TM	Turkmenistan
BF	Burkina Faso	GR	Griechenland			TR	Türkei
BG	Bulgarien	HU	Ungarn	ML	Mali	TT	Trinidad und Tobago
BJ	Benin	IE	Irland	MN	Mongolei	UA	Ukraine
BR	Brasilien	IL	Israel	MR	Mauretanien	UG	Uganda
BY	Belarus	IS	Island	MW	Malawi	US	Vereinigte Staaten von Amerika
CA	Kanada	IT	Italien	MX	Mexiko		
CF	Zentralafrikanische Republik	JP	Japan	NE	Niger	UZ	Usbekistan
CG	Kongo	KE	Kenia	NL	Niederlande	VN	Vietnam
CH	Schweiz	KG	Kirgisistan	NO	Norwegen	YU	Jugoslawien
CI	Côte d'Ivoire	KP	Demokratische Volksrepublik Korea	NZ	Neuseeland	ZW	Zimbabwe
CM	Kamerun			PL	Polen		
CN	China	KR	Republik Korea	PT	Portugal		
CU	Kuba	KZ	Kasachstan	RO	Rumänien		
CZ	Tschechische Republik	LC	St. Lucia	RU	Russische Föderation		
DE	Deutschland	LI	Liechtenstein	SD	Sudan		
DK	Dänemark	LK	Sri Lanka	SE	Schweden		
EE	Estland	LR	Liberia	SG	Singapur		

SDOCID: <WO_____9810352A1_I_>

Beschreibung

Verfahren und Vorrichtung zur rechnergestützten Generierung mindestens eines künstlichen Trainingsdatenvektors für ein neuronales Netz

1. Technischer Hintergrund

Neuronale Netze lernen mit Hilfe von Trainingsdaten. Die Trainingsdaten sind in vielen Anwendungsgebieten sehr verrauscht, beispielsweise bei der Modellierung von Finanzdaten wie Aktien- oder Währungskursen. Somit erhalten die Trainingsdaten zufällige Störungen, die nichts mit der eigentlich zu modellierenden Systemdynamik zu tun haben.

Durch die Approximationsfähigkeit der neuronalen Netze kann aber auch die transiente Struktur des zufälligen Rauschens gelernt werden. Dieses Phänomen wird als Übertrainieren des neuronalen Netzes bezeichnet. Durch ein übertrainiertes neuronales Netz wird in sehr verrauschten Systemen der Lernvorgang des neuronalen Netzes erheblich behindert, da die Verallgemeinerungsfähigkeit des neuronalen Netzes negativ beeinflußt wird.

Dieses Problem gewinnt in Anwendungsgebieten an Bedeutung, in denen nur eine geringe Anzahl von Trainingsdatenvektoren zur Adaption des neuronalen Netzes an die Anwendung, d.h. die zu modellierende, durch die Trainingsdatenvektoren repräsentierte Funktion, verfügbar ist.

Vor allen in diesen Anwendungsgebieten, aber auch allgemein in einem Trainingsverfahren eines neuronalen Netzes ist es vorteilhaft, zusätzliche Trainingsdatenvektoren künstlich zu

SDOCID: <WO_____9810352A1_I_>

generieren, um somit eine größere Trainingsdatenmenge zu erhalten.

2. Stand der Technik

Es ist bekannt, die Generierung der künstlichen Trainingsdatenvektoren durch Verrauschen der verfügbaren Trainingsdatenvektoren der Trainingsdatenmenge durchzuführen. Es ist in diesem Zusammenhang aus dem Dokument [1] bekannt, die Trainingsdatenmenge mit gaußförmigem Rauschen mit dem Mittelwert 0 und einer Varianz σ, die für alle Eingänge des neuronalen Netzes auf den gleichen Wert gesetzt wird, zu ermitteln.

Aus [4] ist es bekannt, Trainingsdaten durch Einführung zusätzlichen Rauschens zu generieren. Dabei ist es bekannt, die sogenannte Jackknife-Prozedur einzusetzen.

Dieses Verfahren weist jedoch einige Nachteile auf.

Dadurch, daß zur Generierung der zusätzlichen Trainingsdatenvektoren als statistische Verteilung, die zur Generierung verwendet wird, ein gaußförmiges Rauschen mit einer Varianz verwendet wird, die für alle Eingänge des neuronalen Netzes auf den gleichen Wert gesetzt wird, werden Trainingsdatenvektoren neu generiert, die keinerlei Aussage über das zu modellierende System enthalten. Die Trainingsdatenvektoren enthalten ferner keinerlei Information über das tatsächliche, dem System zugrundeliegende Rauschen. Somit wird zwar die Trainingsdatenmenge vergrößert, diese muß aber nicht den Lernvorgang unterstützen, da ein fest vorgegebenes Rauschen, das mit der eigentlichen Systemdynamik nichts zu tun hat, zum Trainieren des neuronalen Netzes verwendet wird. Damit kann es dann trotzdem zum Übertrainieren kommen.

Grundlagen über neuronale Netze sind beispielsweise aus dem Dokument [2] bekannt.

Grundlagen über Verwendung der neuronalen Netze in der Ökonomie sind beispielsweise aus dem Dokument [3] bekannt.

3. Kurzbeschreibung der Erfindung

Somit liegt der Erfindung das Problem zugrunde, künstlich neue Trainingsdatenvektoren für ein neuronales Netz zu bilden, wobei jedoch ein Übertrainieren des neuronalen Netzes vermieden wird.

Das Problem wird durch das Verfahren gemäß Patentanspruch 1 sowie durch die Vorrichtung gemäß Patentanspruch 9 gelöst.

Bei dem Verfahren wird nach dem Training des neuronalen Netzes mit verfügbaren Trainingsdatenvektoren einer Trainingsdatenmenge ein Restfehler ermittelt. Aus dem Restfehler wird beispielsweise unter Verwendung eines Gradientenabstiegs-Verfahren ein eingangsbezogener Rückwärtsfehler ermittelt. Die Ermittlung des Rückwärtsfehlers entspricht der üblichen Vorgehensweise während des Trainings eines neuronalen Netzes zur Adaption der einzelnen Gewichte des neuronalen Netzes. Ist der eingangsbezogene Rückwärtsfehler ermittelt worden, wird unter Berücksichtigung des jeweiligen Rückwärtsfehlers eine dem jeweiligen Eingang zugeordnete statistische Verteilung generiert, und der künstliche Trainingsdatenvektor wird unter Berücksichtigung der jeweiligen statistischen Verteilung an den Eingängen des neuronalen Netzes erzeugt.

Mit diesem Verfahren ist es möglich, zusätzliche Trainingsdatenvektoren zu generieren, die eine Information über das neuronale Netz und die Struktur des neuronalen Netzes aktuell nach dem Training des neuronalen Netzes mit den verfügbaren Traindingsdatenvektoren enthalten.

Dies führt dazu, daß die künstlich erzeugten Trainingsdatenvektoren von dem Rückwärtsfehler, der nach dem Training des neuronalen Netzes noch existiert, abhängig sind und somit von

der Performanz des neuronalen Netzes abhängen. Dies führt dazu, daß die erzeugten Trainingsdatenvektoren Information über das zu modellierende System enthalten. Auf diese Weise kann ein Übertrainieren des neuronalen Netzes durch die zusätzlichen Trainingsdatenvektoren vermieden werden.

Die Vorrichtung gemäß Patentanspruch 9 weist eine Recheneinheit auf, die derart eingerichtet ist, daß die oben beschriebenen Verfahrensschritte durchgeführt werden.

Vorteilhafte Weiterbildungen der Erfindung ergeben sich aus den abhängigen Ansprüchen.

Es ist vorteilhaft, bei einer Online-Approximation des neuronalen Netzes, die auch als Online-Training bezeichnet wird, auch die jeweilige statistische Verteilung dem veränderten Trainingsdatensatz anzupassen. Dadurch wird das zu modellierende System noch genauer durch das neuronale Netz modelliert.

4. Kurze Figurenbeschreibung

In den Figuren ist ein Ausführungsbeispiel der Erfindung dargestellt, welches im weiteren näher erläutert wird.

Es zeigen

Fig. 1 ein Ablaufdiagramm, in dem die einzelnen Verfahrensschritte des Verfahrens dargestellt sind;

Fig. 2 eine Skizze, in der eine Rechneranordnung dargestellt ist, mit der das Verfahren durchgeführt werden kann.

5. Figurenbeschreibung

In Fig. 1 sind die einzelnen Verfahrensschritte des Verfahrens dargestellt.

In einem ersten Schritt (101) wird nach abgeschlossenem Training des neuronalen Netzes NN mit einer beliebigen Anzahl von Trainingsdatenvektoren TDVi einer Trainingsdatenmenge TDM von einem Rechner R ein Restfehler RE ermittelt.

Der Restfehler RE wird beispielsweise bei m Trainingsdatenvektoren TDV mit den Trainingsdatenvektoren TDV zugeordneten Sollwerten t_i, wobei mit einem Index i jeweils der Trainingsdatenvektor TDVi eindeutig gekennzeichnet wird, auf folgende Weise ermittelt:

$$RE = \sum_{i=1}^{m} (t_i - y_i)^2 \qquad (1).$$

Es kann jedoch selbstverständliche jede beliebige andere Vorschrift zur Bildung des Restfehlers RE nach dem Training des neuronalen Netzes NN verwendet werden.

Nach Ermittlung des Restfehlers wird ein Rückwärtsfehler RFj ermittelt (102). Die Bestimmung des Rückwärtsfehlers RF kann beispielsweise unter Verwendung eines Gradientenabstiegs-Verfahrens nach den Eingangssignalen der neuronalen Netze NN erfolgen.

Im weiteren wird folgende Nomenklatur für den Trainingsdatenvektor TDVi, sowie die Eingangsdaten des neuronalen Netzes NN, die sich aus den Trainingsdatenvektoren TDVi und einem Rauschtermvektor $\underline{e}$ ergeben, verwendet:

Trainingsdatenvektor TDVi: $\underline{x} = (x_1, \ldots, x_j, \ldots, x_n)$

Rauschtermvektor: $\underline{e} = (e_1, \ldots, e_j, \ldots, e_n)$

Eingangsdatenvektor: $\underline{z} = (x_1+e_1, \ldots, x_j+e_j, \ldots, x_n+e_n)$

Mit n wird eine Anzahl von Komponenten der Trainingsdatenvektoren TDVi, der Rauschtermvektoren $\underline{e}$ bzw. der Eingangsdatenvektoren $\underline{z}$ bezeichnet.

Der Rückwärtsfehler RFj wird für jeden Eingang j des neuronalen Netzes NN individuell ermittelt, wobei mit einem Index j jeweils eine Komponente des Eingangsdatenvektors $\underline{z}$ bzw. ein Eingang des neuronalen Netzes NN eindeutig bezeichnet wird,.

Dies erfolgt beispielsweise nach dem bekannten Gradientenabstiegs-Verfahren zum Trainieren des neuronalen Netzes NN. Somit ergibt sich der Rückwärtsfehler RF des Eingangsdatenvektors $\underline{z}$ aus den jeweiligen partiellen Ableitungen des Restfehlers RE nach den einzelnen Eingangssignalen z_j.

$$RFj = \frac{\partial E^t}{\partial z_j^t} \qquad (2).$$

Mit dem Symbol t wird jeweils ein Zeitpunkt eindeutig gekennzeichnet, zu dem die Ermittlung des Rückwärtsfehlers RFj erfolgt.

Unter Berücksichtigung des Rückwärtsfehlers RFj wird nun in einem dritten Schritt (103) eine statistische Verteilung S_j für jeden Eingang j des neuronalen Netzes NN gebildet.

Als statistische Verteilung S_j kann allgemein jede beliebige statistische Verteilung zur Beschreibung eines Zufallsprozesses verwendet werden.

Im folgenden werden zur einfacheren Darstellung jedoch nur eine uniforme statistische Verteilung und eine gaußförmige statistische Verteilung S_j näher erläutert.

Für die uniforme Verteilung ergibt sich eine uniforme Verteilungsbreite s_j^t beispielsweise nach folgender Vorschrift:

$$s_j^t = \frac{1}{t} \sum_{k=1}^{t} \left| \frac{\partial E^k}{\partial z_j^k} \right| \qquad (3).$$

Der Rauschterm $e_j{}^t$, der zur Bildung mindestens eines künstlichen Trainingsdatenvektors KTDV verwendet wird, liegt im Bereich des Intervalls:

$$e_j^t \in \left[-s_j^t , s_j^t \right] \qquad (4).$$

Bei einer gaußförmigen Verteilung ergibt sich die Rauschbreite s_j^t nach der folgenden Vorschrift:

$$s_j^t = \frac{1}{t} \sum_{k=1}^{t} \left(\frac{\partial E^k}{\partial z_j^k} \right)^2 \qquad (5).$$

Der Rauschterm $e_j{}^t$ ergibt sich für diesen Beispielsfall nach folgender Vorschrift:

$$e_j^t \in N\left(0, \sqrt{s_j^t} \right) \qquad (6).$$

Mit $N\left(0, \sqrt{s_j^t} \right)$ wird ein normal verteilter Zufallsprozeß mit einem Mittelwert 0 und der Standardabweichung $\sqrt{s_j^t}$ bezeichnet.

Unter Verwendung der jeweiligen statistischen Verteilung S_j wird der mindestens eine künstliche Trainingsdatenvektor KTDV generiert (104).

ISDOCID: <WO_____9810352A1_I_>

Anschaulich bedeutet dies, daß der künstliche Trainingsdatenvektor KTDV durch den jeweils durch die statistische Verteilung S_j beschriebenen Zufallsprozeß generiert wird.

Durch diese Vorgehensweise haben die künstlichen Trainingsdatenvektoren KTDV statistisch die gleichen Eigenschaften wie die original verfügbaren Trainingsdatenvektoren TDVi.

Dies kann durch Verrauschen mit einem einheitlichen Rauschprozeß, wie dieser aus dem Dokument [1] bekannt ist, nicht erreicht werden.

Anschaulich läßt sich das Verfahren folgendermaßen beschreiben. Die Eingangsdaten, repräsentiert durch die original verfügbaren Trainingsdatenvektoren TDVi, werden mit einem Zufallsprozeß verrauscht, der für jeden Eingang abhängig ist von der jeweiligen Gradienten-Fehlerfunktion nach den Eingangssignalen.

Die jeweilige Rauschbreite s_j^t kann anschaulich folgendermaßen interpretiert werden:

- Kleine Werte von s_j^t werden durch Eingänge erzeugt, die eine gute Abbildung auf die Trainingsdaten erzeugen.

- Kleine Werte der Rauschbreite s_j^t können aber auch auf unwichtige, für das Training des neuronalen Netzes NN nicht relevante Eingänge hinweisen.

In beiden Fällen ist es im Sinne der Aufgabe, daß kein oder nur wenig Rauschen zu den ursprünglich vorhandenen Trainingsdatenvektoren TDVi hinzugefügt wird.

- Große Werte der Rauschbreite s_j^t deuten darauf hin, daß der Eingang j erhebliche Bedeutung aufweist, aber das neuronale

SDOCID: <WO_____9810352A1_I_>

Netz NN bisher nicht in der Lage ist, eine adäquate Abbildung zu lernen.

- Rauschen auf Eingängen mit einer großen Rauschbreite s_j^t "versteift" das neuronale Netz NN, wodurch ein Übertrainieren verhindert wird, was zu einer besseren Verallgemeinerungsfähigkeit des neuronalen Netzes NN führt.

Dieses Verfahren kann besonders vorteilhaft in Anwendungssituationen verwendet werden, in denen nur eine relativ geringe Anzahl von Trainingsdatenvektoren TDVi zum Trainieren des neuronalen Netzes NN verfügbar ist.

Durch das Verfahren kann die Trainingsdatenmenge TDM künstlich erheblich erweitert werden zu einer erweiterten Trainingsdatenmenge KTDM, ohne die eigentliche Systemdynamik zu verfälschen, da die statistischen Eigenschaften der Trainingsdatenmenge TDM auch in den künstlichen Trainingsdatenvektoren KTDV enthalten sind.

Eine typische Anwendungssituation des Verfahrens liegt in der Analyse von Finanzmärkten, beispielsweise von Aktienmärkten oder auch Wertpapiermärkten.

Wird zum Trainieren des neuronalen Netzes NN ein sog. Online-Lernverfahren eingesetzt, bei dem während der Generalisierungsphase des neuronalen Netzes NN weitere Trainingsdatenvektoren TDV ermittelt werden, mit denen das neuronale Netz NN in der Generalisierungsphase weiter trainiert wird, so ist es in einer Weiterbildung des Verfahrens vorgesehen, abhängig von der Änderung der Trainingsdatenmenge TDM, die durch die Trainingsdatenvektoren TDVi gebildet wird, auch die statistischen Verteilungen S_j anzupassen. Dies kann beispielsweise für eine uniforme Verteilung nach folgender Vorschrift erfolgen:

$$s_j^{t+1} = (1 - \alpha) \cdot s_j^t + \alpha \cdot \left| \frac{\partial E^k}{\partial z_j^k} \right| \qquad (7).$$

Für eine gaußförmige statistische Verteilung S_j kann beispielsweise folgende Adaptionsvorschrift eingesetzt werden:

$$\left(s_j^{t+1}\right)^2 = (1 - \alpha) \cdot \left(s_j^t\right)^2 + \alpha \cdot \left(\frac{\partial E^k}{\partial z_j^k} \right)^2 \qquad (8).$$

Mit α wird ein Adaptionsfaktor bezeichnet.

Es hat sich als vorteilhaft herausgestellt, als Wert für den Adaptionsfaktor α den Kehrwert der Anzahl der Trainingsdatenvektoren TDVi zu verwenden.

Im Rahmen dieses Dokumentes wurden folgende Veröffentlichungen zitiert:

[1] C. M. Bishop, Neural Networks for Pattern Recognition, Clarendon Press, Oxford, U. K., ISBN 0-19-853864-2, S. 346 - 349, 1994

[2] G. Deco und D: Obradovic, An Information-Theoretic Approach to Neural Computing, Springer Verlag, New York, USA, ISBN 0-387-94666-7, S. 23 - 37, 1996

[3] H. Zimmermann und Rehkugler, Neuronale Netze in der Ökonomie, Kapitel 1, Neuronale Netze als Entscheidungskalkül, Vahlen Verlag, München, ISBN 3-8006-1871-0, S. 3 - 20, 1994

[4] L. Holmström und P. Koistinen, Using Additive Noise in Back-Propagation Training, IEEE Transactions on Neural Networks, Vol. 3, No. 1, S. 24 - 38, Januar 1992

3DOCID: <WO_____9810352A1_I_>

Patentansprüche

1. Verfahren zur rechnergestützten Generierung mindestens eines künstlichen Trainingsdatenvektors für ein Neuronales Netz (NN),
- bei dem ein Restfehler (RE) nach einem Training des Neuronalen Netzes (NN) mit Trainingsdatenvektoren (TDVi) einer Trainingsdatenmenge (TDM) ermittelt wird (101),
- bei dem aus dem Restfehler (RE) für mindestens einen Eingang (j) des Neuronalen Netzes (NN) ein Rückwärtsfehler (RFj) ermittelt wird (102),
- bei dem jeweils eine dem Eingang (j) zugeordnete statistische Verteilung (S_j) generiert wird, die von der Größe des jeweiligen Rückwärtsfehlers (RFj) abhängig ist (103),
- bei dem der künstliche Trainingsdatenvektor generiert wird unter jeweiliger Verwendung der einem Eingang (j) zugeordneten statistischen Verteilung (S_j) (104).

2. Verfahren nach Anspruch 1,
bei dem der Rückwärtsfehler (RFj) mit einem Gradientenabstiegsverfahren ermittelt wird.

3. Verfahren nach Anspruch 2,
bei dem das Gradientenabstiegsverfahren mit einem Backpropagation-Verfahren realisiert wird.

4. Verfahren nach einem der Ansprüche 1 bis 3,
bei dem die statistische Verteilung (S_j) durch eine uniforme Verteilung gebildet wird, wobei eine Verteilungsbreite der uniformen Verteilung von dem Rückwärtsfehler (RFj) abhängig ist.

5. Verfahren nach einem der Ansprüche 1 bis 3,
bei dem die statistische Verteilung (S_j) durch eine gaußförmige Verteilung gebildet wird, wobei eine Verteilungsbreite der gaußförmigen Verteilung von dem Rückwärtsfehler (RFj) abhängig ist.

ISDOCID: <WO_____9810352A1_I_>

6. Verfahren nach Anspruch 4 oder 5,
bei dem bei einer Online-Approximation des Neuronalen Netzes (NN) nach einer beliebigen Anzahl neuer Trainingsdatenvektoren (TDVi) die statistische Verteilung (S_j) an einen neu ermittelten Rückwärtsfehler angepaßt wird.

7. Verfahren nach einem der Ansprüche 1 bis 6 eingesetzt zur Modellierung einer nichtlinearen Funktion, die mit einer geringen Anzahl von Trainingsdatenvektoren beschrieben wird.

8. Verfahren nach einem der Ansprüche 1 bis 6 eingesetzt in der Analyse eines Finanzmarktes.

9. Vorrichtung zur rechnergestützten Generierung mindestens eines künstlichen Trainingsdatenvektors für ein Neuronales Netz (NN),
mit einer Recheneinheit, die derart eingerichtet ist, daß
- ein Restfehler (RE) nach einem Training des Neuronalen Netzes (NN) mit Trainingsdatenvektoren (TDVi) einer Trainingsdatenmenge (TDM) ermittelt wird (101),
- aus dem Restfehler (RE) für mindestens einen Eingang (j) des Neuronalen Netzes (NN) ein Rückwärtsfehler (RFj) ermittelt wird (102),
- jeweils eine dem Eingang (j) zugeordnete statistische Verteilung (S_j) generiert wird, die von der Größe des jeweiligen Rückwärtsfehlers (RFj) abhängig ist (103),
- der künstliche Trainingsdatenvektor generiert wird unter jeweiliger Verwendung der einem Eingang (j) zugeordneten statistischen Verteilung (S_j) (104).

10. Vorrichtung nach Anspruch 9,
bei der die Recheneinheit derart eingerichtet ist, daß der Rückwärtsfehler (RFj) mit einem Gradientenabstiegsverfahren ermittelt wird.

11. Vorrichtung nach Anspruch 9 oder 10,

SDOCID: <WO_____9810352A1_I_>

bei der die Recheneinheit derart eingerichtet ist, daß die statistische Verteilung (S_j) durch eine uniforme Verteilung gebildet wird, wobei eine Verteilungsbreite der uniformen Verteilung von dem Rückwärtsfehler (RFj) abhängig ist.

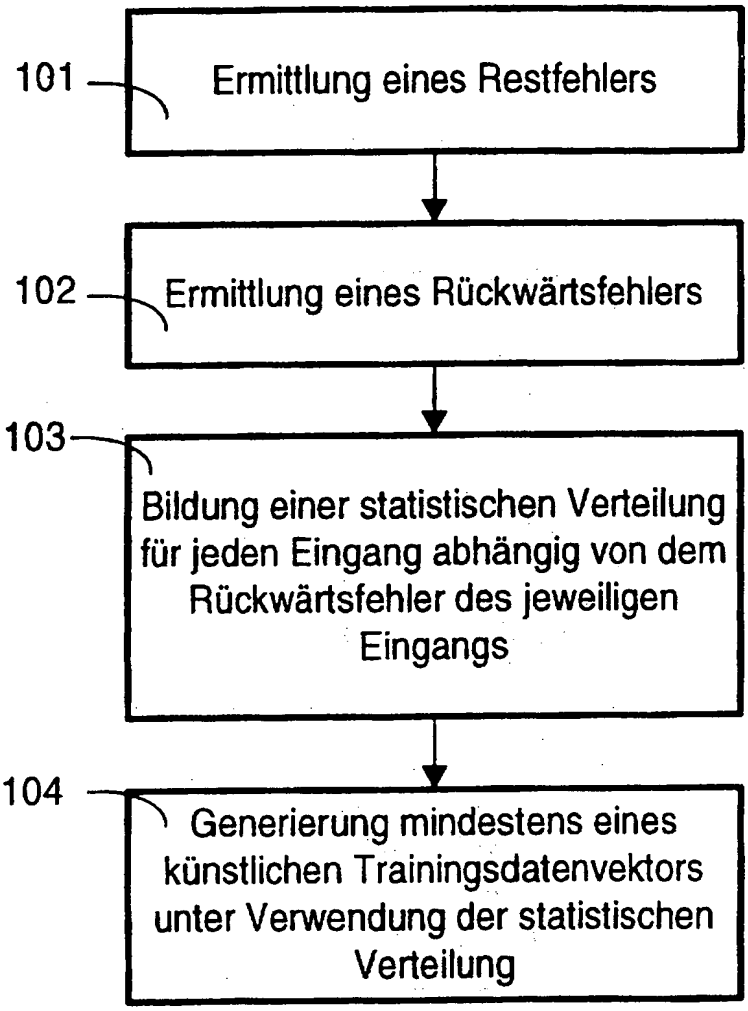
12. Vorrichtung nach einem der Ansprüche 9 bis 11, bei der die Recheneinheit derart eingerichtet ist, daß die statistische Verteilung (S_j) durch eine gaußförmige Verteilung gebildet wird, wobei eine Verteilungsbreite der gaußförmigen Verteilung von dem Rückwärtsfehler (RFj) abhängig ist.

13. Vorrichtung nach Anspruch 11 oder 12, bei der die Recheneinheit derart eingerichtet ist, daß bei einer Online-Approximation des Neuronalen Netzes (NN) nach einer beliebigen Anzahl neuer Trainingsdatenvektoren (TDVi) die statistische Verteilung (S_j) an einen neu ermittelten Rückwärtsfehler angepaßt wird.

14. Vorrichtung nach einem der Ansprüche 9 bis 13 eingesetzt zur Modellierung einer nichtlinearen Funktion, die mit einer geringen Anzahl von Trainingsdatenvektoren beschrieben wird.

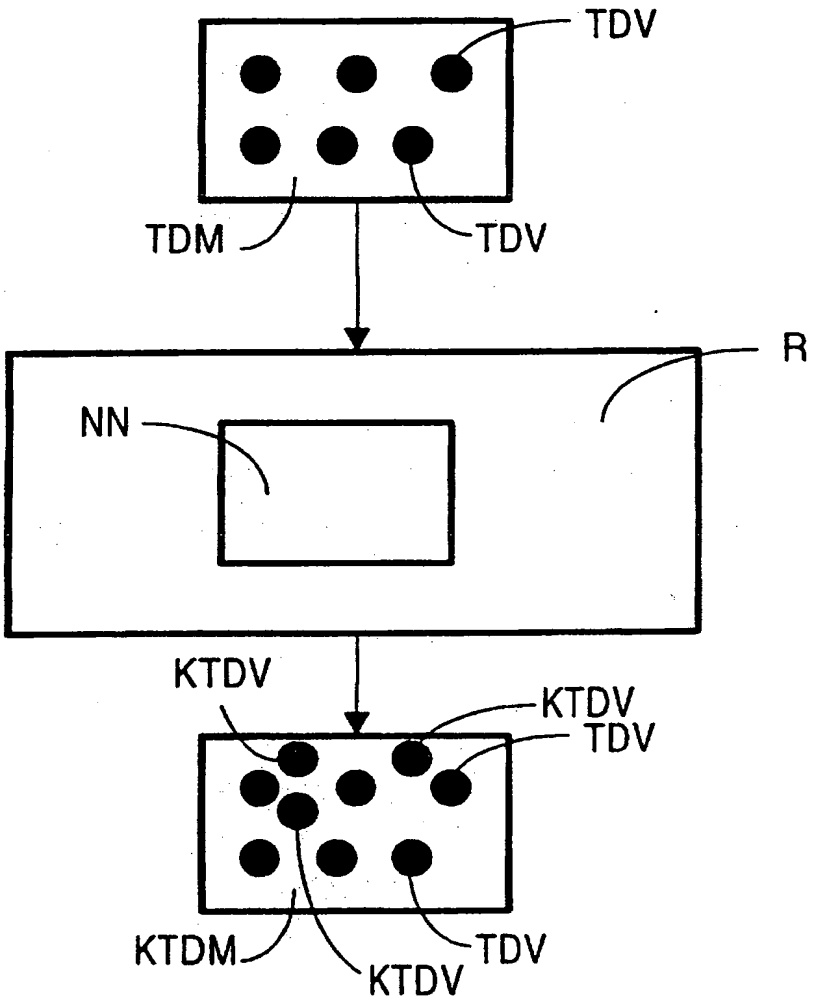
15. Vorrichtung nach einem der Ansprüche 9 bis 13 eingesetzt in der Analyse eines Finanzmarktes.

FIG 1



SDOCID: <WO_____9810352A1_I_>

FIG 2



INTERNATIONAL SEARCH REPORT

International Application No
PCT/DE 97/01497

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 G06F15/80

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 95 11486 A (MILES INC) 27 April 1995 see page 10, line 1 - line 32; figure 3	1,9
A	US 5 359 699 A (TONG DAVID W ET AL) 25 October 1994 see column 4, line 5 - line 42	1,9

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

1

Date of the actual completion of the international search	Date of mailing of the international search report
3 December 1997	16/12/1997

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Schenkels, P

Form PCT/ISA/210 (second sheet) (July 1992)

ISDOCID: <WO_____9810352A1_I_>

INTERNATIONAL SEARCH REPORT

International Application No
PCT/DE 97/01497

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
A	SUNGZOON CHO ET AL: "Evolution of neural network training set through addition of virtual samples" PROCEEDINGS OF 1996 IEEE INTERNATIONAL CONFERENCE ON EVOLUTIONARY COMPUTATION (ICEC'96) (CAT. NO.96TH8114), PROCEEDINGS OF IEEE INTERNATIONAL CONFERENCE ON EVOLUTIONARY COMPUTATION, NAGOYA, JAPAN, 20-22 MAY 1996, ISBN 0-7803-2902-3, 1996, NEW YORK, NY, USA, IEEE, USA, pages 685-688, XP002048950 see page 685, left-hand column, line 1 - page 686, right-hand column, line 30; figures 1,2 ---	1,9
A	DATUM M S ET AL: "An artificial neural network for sound localization using binaural cues" JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA, JULY 1996, ACOUST. SOC. AMERICA THROUGH AIP, USA, vol. 100, no. 1, ISSN 0001-4966, pages 372-383, XP002048951 see abstract see page 372, left-hand column, line 1 - right-hand column, line 31 see page 376, left-hand column, line 26 - page 377, left-hand column, line 29; figures 1,2 ---	1-3,9,10
A	SLICHER A W R ET AL: "An innovative approach to training neural networks for strategic management of construction firms" DEVELOPMENTS IN NEURAL NETWORKS AND EVOLUTIONARY COMPUTING FOR CIVIL AND STRUCTURAL ENGINEERING, DEVELOPMENTS IN NEURAL NETWORKS AND EVOLUTIONARY COMPUTING FOR CIVIL AND STRUCTURAL ENGINEERING, CAMBRIDGE, UK, 28-30 AUG. 1995, ISBN 0-948749-36-9, 1995, EDINBURGH, UK, CIVIL-COMP PRESS, UK, pages 87-93, XP002048952 see abstract see page 89, left-hand column, line 5 - page 92, left-hand column, line 8; figure 1 -----	1,9

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

Int. ational Application No
PCT/DE 97/01497

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9511486 A	27-04-95	US 5444796 A	22-08-95
		AU 7981794 A	08-05-95
		CA 2174522 A	27-04-95
		EP 0724750 A	07-08-96
		US 5590218 A	31-12-96
US 5359699 A	25-10-94	NONE	

NSDOCID: <WO_____9810352A1_I_>

INTERNATIONALER RECHERCHENBERICHT

Internationales Aktenzeichen
PCT/DE 97/01497

A. KLASSIFIZIERUNG DES ANMELDUNGSGEGENSTANDES
IPK 6 G06F15/80

Nach der Internationalen Patentklassifikation (IPK) oder nach der nationalen Klassifikation und der IPK

B. RECHERCHIERTE GEBIETE

Recherchierter Mindestprüfstoff (Klassifikationssystem und Klassifikationssymbole)
IPK 6 G06F

Recherchierte aber nicht zum Mindestprüfstoff gehörende Veröffentlichungen, soweit diese unter die recherchierten Gebiete fallen

Während der internationalen Recherche konsultierte elektronische Datenbank (Name der Datenbank und evtl. verwendete Suchbegriffe)

C. ALS WESENTLICH ANGESEHENE UNTERLAGEN

Kategorie*	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
A	WO 95 11486 A (MILES INC) 27.April 1995 siehe Seite 10, Zeile 1 - Zeile 32; Abbildung 3 ---	1,9
A	US 5 359 699 A (TONG DAVID W ET AL) 25.Oktober 1994 siehe Spalte 4, Zeile 5 - Zeile 42 --- -/--	1,9

[X] Weitere Veröffentlichungen sind der Fortsetzung von Feld C zu entnehmen

[X] Siehe Anhang Patentfamilie

* Besondere Kategorien von angegebenen Veröffentlichungen :
"A" Veröffentlichung, die den allgemeinen Stand der Technik definiert, aber nicht als besonders bedeutsam anzusehen ist
"E" älteres Dokument, das jedoch erst am oder nach dem internationalen Anmeldedatum veröffentlicht worden ist
"L" Veröffentlichung, die geeignet ist, einen Prioritätsanspruch zweifelhaft erscheinen zu lassen, oder durch die das Veröffentlichungsdatum einer anderen im Recherchenbericht genannten Veröffentlichung belegt werden soll oder die aus einem anderen besonderen Grund angegeben ist (wie ausgeführt)
"O" Veröffentlichung, die sich auf eine mündliche Offenbarung, eine Benutzung, eine Ausstellung oder andere Maßnahmen bezieht
"P" Veröffentlichung, die vor dem internationalen Anmeldedatum, aber nach dem beanspruchten Prioritätsdatum veröffentlicht worden ist
"T" Spätere Veröffentlichung, die nach dem internationalen Anmeldedatum oder dem Prioritätsdatum veröffentlicht worden ist und mit der Anmeldung nicht kollidiert, sondern nur zum Verständnis des der Erfindung zugrundeliegenden Prinzips oder der ihr zugrundeliegenden Theorie angegeben ist
"X" Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann allein aufgrund dieser Veröffentlichung nicht als neu oder auf erfinderischer Tätigkeit beruhend betrachtet werden
"Y" Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann nicht als auf erfinderischer Tätigkeit beruhend betrachtet werden, wenn die Veröffentlichung mit einer oder mehreren anderen Veröffentlichungen dieser Kategorie in Verbindung gebracht wird und diese Verbindung für einen Fachmann naheliegend ist
"&" Veröffentlichung, die Mitglied derselben Patentfamilie ist

1

Datum des Abschlusses der internationalen Recherche	Absendedatum des internationalen Recherchenberichts
3.Dezember 1997	16/12/1997

Name und Postanschrift der Internationalen Recherchenbehörde	Bevollmächtigter Bediensteter
Europäisches Patentamt, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31–70) 340–2040, Tx. 31 651 epo nl, Fax: (+31–70) 340–3016	Schenkels, P

Formblatt PCT/ISA/210 (Blatt 2) (Juli 1992)

SDOCID: <WO_____9810352A1_I_>

INTERNATIONALER RECHERCHENBERICHT

Internationales Aktenzeichen
PCT/DE 97/01497

C.(Fortsetzung) ALS WESENTLICH ANGESEHENE UNTERLAGEN

Kategorie	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
A	SUNGZOON CHO ET AL: "Evolution of neural network training set through addition of virtual samples" PROCEEDINGS OF 1996 IEEE INTERNATIONAL CONFERENCE ON EVOLUTIONARY COMPUTATION (ICEC'96) (CAT. NO.96TH8114), PROCEEDINGS OF IEEE INTERNATIONAL CONFERENCE ON EVOLUTIONARY COMPUTATION, NAGOYA, JAPAN, 20-22 MAY 1996, ISBN 0-7803-2902-3, 1996, NEW YORK, NY, USA, IEEE, USA, Seiten 685-688, XP002048950 siehe Seite 685, linke Spalte, Zeile 1 - Seite 686, rechte Spalte, Zeile 30; Abbildungen 1,2 ---	1,9
A	DATUM M S ET AL: "An artificial neural network for sound localization using binaural cues" JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA, JULY 1996, ACOUST. SOC. AMERICA THROUGH AIP, USA, Bd. 100, Nr. 1, ISSN 0001-4966, Seiten 372-383, XP002048951 siehe Zusammenfassung siehe Seite 372, linke Spalte, Zeile 1 - rechte Spalte, Zeile 31 siehe Seite 376, linke Spalte, Zeile 26 - Seite 377, linke Spalte, Zeile 29; Abbildungen 1,2 ---	1-3,9,10
A	SLICHER A W R ET AL: "An innovative approach to training neural networks for strategic management of construction firms" DEVELOPMENTS IN NEURAL NETWORKS AND EVOLUTIONARY COMPUTING FOR CIVIL AND STRUCTURAL ENGINEERING, DEVELOPMENTS IN NEURAL NETWORKS AND EVOLUTIONARY COMPUTING FOR CIVIL AND STRUCTURAL ENGINEERING, CAMBRIDGE, UK, 28-30 AUG. 1995, ISBN 0-948749-36-9, 1995, EDINBURGH, UK, CIVIL-COMP PRESS, UK, Seiten 87-93, XP002048952 siehe Zusammenfassung siehe Seite 89, linke Spalte, Zeile 5 - Seite 92, linke Spalte, Zeile 8; Abbildung 1 -----	1,9

1

INTERNATIONALER RECHERCHENBERICHT

Angaben zu Veröffentlichungen, die zur selben Patentfamilie gehören

Internationales Aktenzeichen
PCT/DE 97/01497

Im Recherchenbericht angeführtes Patentdokument	Datum der Veröffentlichung	Mitglied(er) der Patentfamilie	Datum der Veröffentlichung
WO 9511486 A	27-04-95	US 5444796 A	22-08-95
		AU 7981794 A	08-05-95
		CA 2174522 A	27-04-95
		EP 0724750 A	07-08-96
		US 5590218 A	31-12-96
US 5359699 A	25-10-94	KEINE	

SDOCID: <WO_____9810352A1_I_>